

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

IN RE APPLICATION OF: Eva KONDOROSI, et al.

SERIAL NO.: NEW U.S. PCT APPLICATION

FILED: HERewith

INTERNATIONAL APPLICATION NO.: PCT/FR99/01342

INTERNATIONAL FILING DATE: 08 JUNE 1999

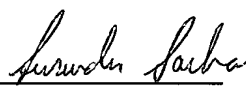
FOR: PLANT PROTEIN WITH REPEATED WD40 MOTIFS, NUCLEIC ACID CODING  
FOR SAID PROTEIN, AND USES THEREOF**REQUEST FOR PRIORITY UNDER 35 U.S.C. 119**  
**AND THE INTERNATIONAL CONVENTION**Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

In the matter of the above-identified application for patent, notice is hereby given that  
the applicant claims as priority:

<u>COUNTRY</u>	<u>APPLICATION NO.</u>	<u>DAY/MONTH/YEAR</u>
FRANCE	98/07174	08 JUNE 1998

A certified copy of the corresponding Convention application(s) was submitted to the  
International Bureau in PCT Application No. PCT/FR99/01342. **Receipt of the certified  
copy(s) by the International Bureau in a timely manner under PCT Rule 17.1(a) has  
been acknowledged as evidenced by the attached PCT/IB/304.**

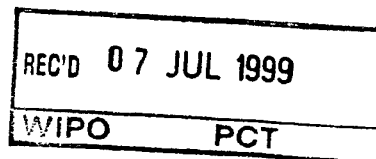
Respectfully submitted,  
OBLON, SPIVAK, McCLELLAND,  
MAIER & NEUSTADT, P.C.**22850**  
Norman F. Oblon  
Attorney of Record  
Registration No. 24,618  
Surinder Sachar  
Registration No. 34,423





PCT/FR 99/01342

FR 99/01342



# BREVET D'INVENTION

CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION

COPIE OFFICIELLE

**PRIORITY  
DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH RULE 17.1(a) OR (b)

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le 29 JUIN 1999

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef du Département des brevets

Martine PLANCHE

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS Cédex 08  
Téléphone : 01 53 04 53 04  
Télécopie : 01 42 93 59 30



26 bis, rue de Saint Pétersbourg  
75800 Paris Cedex 08  
Téléphone : 01 53 04 53 04 Télécopie : 01 42 93 59 30

**REQUÊTE EN DÉLIVRANCE**

Confirmation d'un dépôt par télécopie ☐

Cet imprimé est à remplir à l'encre noire en lettres capitales

Réserve à l'INPI

DATE DE REMISE DES PIÈCES

08 JUIN 1998

N° D'ENREGISTREMENT NATIONAL

98 07174 -

DÉPARTEMENT DE DÉPÔT

DATE DE DÉPÔT

08 JUIN 1998

1 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE  
À QUI LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE

CABINET ORES  
6 avenue de Messine  
75008 PARIS  
FRANCE

n° du pouvoir permanent références du correspondant téléphone  
MJPsdt644/39 01.45.62.75.00

2 DEMANDE Nature du titre de propriété industrielle

☒ brevet d'invention ☐ demande divisionnaire

☐ certificat d'utilité ☐ transformation d'une demande de brevet européen

demande initiale

☐ brevet d'invention

☐ certificat d'utilité n°

date

Établissement du rapport de recherche

☐ différé ☒ immédiat

Le demandeur, personne physique, requiert le paiement échelonné de la redevance

☐ oui ☐ non

Titre de l'invention (200 caractères maximum)

PROTEINE VEGETALE A MOTIFS WD40 REPETES, ACIDE NUCLEIQUE CODANT POUR LADITE PROTEINE, ET LEURS APPLICATIONS

3 DEMANDEUR (S)

n° SIREN

code APE-NAF

Nom et prénoms (souligner le nom patronymique) ou dénomination

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE (CNRS)

Forme juridique

Nationalité (s) FRANCAISE

Adresse (s) complète (s)

3 rue Michel Ange  
75794 PARIS CEDEX 16

Pays

FRANCE

4 INVENTEUR (S) Les inventeurs sont les demandeurs

☐ oui ☒ non Si la réponse est non, fournir une désignation séparée

5 RÉDUCTION DU TAUX DES REDEVANCES

☐ requise pour la 1ère fois ☐ requise antérieurement au dépôt : joindre copie de la décision d'admission

6 DÉCLARATION DE PRIORITÉ OU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTÉRIEURE

pays d'origine

numéro

date de dépôt

nature de la demande

7 DIVISIONS antérieures à la présente demande n°

date

n°

date

8 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE  
(nom et qualité du signataire - n° d'inscription)

SIGNATURE DU PRÉPOSÉ À LA RÉCEPTION

SIGNATURE APRÈS ENREGISTREMENT DE LA DEMANDE À L'INPI

Béatrice ORES (N° 92-4046)

DÉSIGNATION DE L'INVENTEUR

(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

DIVISION ADMINISTRATIVE DES BREVETS

MJPsd644/39

26bis, rue de Saint-Petersbourg

75800 Paris Cédex 08

Tél. : 01 53 04 53 04 - Télécopie : 01 42 93 59 30

N° D'ENREGISTREMENT NATIONAL

9807174

TITRE DE L'INVENTION :

PROTEINE VEGETALE A MOTIFS WD40 REPETES, ACIDE NUCLEIQUE CODANT POUR  
LADITE PROTEINE, ET LEURS APPLICATIONS

LE(S) SOUSSIGNÉ(S)

CABINET ORES  
6 avenue de Messine  
75008 PARIS  
FRANCE


DÉSIGNE(NT) EN TANT QU'INVENTEUR(S) (indiquer nom, prénoms, adresse et souligner le nom patronymique) :

- KONDOROSI Eva  
10 allée de la Dame Alips  
91190 GIF SUR YVETTE  
FRANCE
- CEBOLLA Angel  
15 rue Juliette Adam  
91190 GIF SUR YVETTE  
FRANCE
- KONDOROSI Adam  
10 allée de la Dame Alips  
91190 GIF SUR YVETTE  
FRANCE

NOTA : A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient (société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire

Le 8 juillet 1998



Béatrice ORES (N° 92-4046)

L'invention est relative au clonage de gènes intervenant dans la régulation de la division cellulaire chez les végétaux, et à leurs utilisations.

La plupart des organes végétaux se développent  
 5 après la germination, par différenciation à partir des  
 méristèmes. Préalablement à la différenciation, se  
 produisent dans les méristèmes le ralentissement puis  
 l'arrêt du cycle de division cellulaire. Simultanément,  
 on observe fréquemment une augmentation de la taille des  
 10 cellules, et une répllication du génome non accompagnée de  
 mitose, dénommée « endorépllication ». L'endorépllication  
 est un phénomène bien connu lors du développement de  
 tissus de réserve ; KOWLES [Genome, 35, pp. 68-77,  
 (1992)] mentionne ainsi une ploïdie de 6C à 384C lors du  
 15 développement de l'endosperme chez le maïs.

Les phénomènes intervenant lors de l'arrêt de  
 la division cellulaire précédant la différenciation  
 jouent un rôle essentiel dans le développement et  
 l'ontogénèse végétale. Les mécanismes impliqués dans ces  
 20 phénomènes sont encore mal connus ; il semble que  
 l'inhibition du facteur promoteur de la phase M, et  
 l'induction des protéine-kinases de la phase S (GRAFI,  
 Science, 269, pp. 1262-1264, (1995)] seraient impliqués.  
 Toutefois, on n'a jusqu'à présent identifié aucun facteur  
 25 directement impliqué dans ce mécanisme chez les plantes.

Les Inventeurs ont entrepris l'étude de ce  
 mécanisme, dans le but de découvrir des moyens de le  
 contrôler, et d'agir par son intermédiaire sur le  
 développement et l'ontogénèse végétale.

30 Ils ont choisi comme modèle d'étude le système  
 symbiotique *Rhizobium*/légumineuses. Dans ce système, les  
 facteurs Nod de nature lipooligosaccharidique, produits  
 par les *Rhizobium*, constituent des signaux mitogènes qui  
 induisent localement la formation d'un nouveau méristème,  
 35 à partir duquel se différencient les cellules formant les  
 nodosités racinaires [TRUCHET, Nature, 351, pp. 670-673,

(1991) ; YANG, Plant Cell, 6, pp. 1415-1426, (1994) ; SAVOURE, EMBO.J., 13, pp. 1093-1102, (1994)]. Les nodosités comprennent 3 zones principales : une zone apicale, constituée de cellules méristématiques ; une  
 5 zone intermédiaire d'invasion, ou de différenciation (zone II), où intervient l'infection des cellules par les bactéries, ainsi que l'arrêt de la division cellulaire, accompagné d'endoréplication et d'augmentation de la taille des cellules, et suivi par leur différenciation ;  
 10 et une zone de fixation (zone III), formée de cellules différenciées infectées par les bactéries, et où intervient la fixation de l'azote.

Au cours de cette étude, les Inventeurs ont isolé, à partir de nodosités de luzerne (*Medicago sativa*), un gène jouant un rôle essentiel dans l'arrêt du  
 15 cycle cellulaire et l'induction de l'endoréplication.

Ce gène, dénommé ci-après *ccc52Ms*, code pour un polypeptide de 475 acides aminés, ayant une masse moléculaire théorique de 52 kDa, dénommé ci-après  
 20 *ccc52Ms* ; cette protéine comprend des motifs WD répétés ; la présence de tels motifs implique que *ccc52Ms* appartient à la superfamille des protéines à motifs répétés. Les motifs WD répétés comprennent environ 40 aminoacides contenant un certain nombre d'acides aminés  
 25 conservés dont le motif WD (Trp-Asp) qui se situe fréquemment à une extrémité du motif répété [NEER et al., Nature, 371, pp. 297-300, (1994)]. Les membres de cette famille régulent différentes fonctions, telles que la transduction de signal, la transcription, l'épissage de  
 30 pré-ARNm, l'organisation du cytosquelette, la fusion vésiculaire ou le cycle cellulaire. Bien que la structure générale soit globalement similaire dans toutes les protéines, la grande variété fonctionnelle des motifs WD répétés suggère que ces motifs répétés se sont  
 35 différenciés et sont devenus fonctionnellement spécialisés. Une homologie fonctionnelle se reflète dans



le nombre de motifs WD répétés, par une homologie importante des motifs WD répétés à des positions équivalentes dans différentes protéines, par rapport à d'autres motifs répétés dans les mêmes protéines, et par  
 5 une similarité significative des extrémités C et N terminales.

La comparaison de la séquence de ccc52Ms avec les séquences de protéines connues a fait apparaître une homologie élevée avec des protéines à motifs WD40 répétés  
 10 qui interviennent dans la régulation du cycle cellulaire, et, plus spécifiquement, avec les protéines fzf de *Drosophile* (57% d'identité), hct1 de *Saccharomyces cerevisiae* (46% d'identité), et srw1 de *Schizosaccharomyces pombe* (52% d'identité), qui  
 15 appartiennent à la famille « fizzy-related ».

Les recherches effectuées par les Inventeurs dans des bases de données de séquences ont en outre fait apparaître une homologie importante avec le produit du gène fzf de *X. laevis* ou de *C. elegans*.

En ce qui concerne plus particulièrement la sous-famille fzf, il a été montré chez la *Drosophile* que le gène fzf est exprimé en fin de prolifération cellulaire pendant l'embryogénèse. Le produit de ce gène entraîne une diminution des cyclines mitotiques, et est  
 20 nécessaire pour l'arrêt de la prolifération cellulaire et le début des endocycles [SIGRIST et LEHNER, Cell, 90, pp. 671-681, (1997)]. Chez *Saccharomyces cerevisiae*, hct1 est nécessaire pour la protéolyse de la cycline mitotique, Clb2 [SCHWAB et al., Cell, 90, pp. 683-693,  
 25 (1997)]. Chez *Schizosaccharomyces pombe*, le produit du gène srw1 contrôle le cycle cellulaire et la différenciation en régulant négativement les complexes Cdc2/cycline de type B [YAMAGUCHI et al., Mol. Biol. Cell., 8, 2475-2486, (1997)].  
 30

Chez les plantes, aucune protéine de la famille fzf n'a été décrite jusqu'à présent.  
 35

Aucune autre séquence peptidique végétale présentant un degré d'homologie aussi important n'a été retrouvé dans les banques de données. L'existence d'un gène codant pour une protéine à motifs WD40 répétés et son isolement à partir d'ADNc de carotte ont été récemment décrits [LUO et al., Plant Mol. Biol., 34, pp. 325-330, (1997)]. Cependant, le produit de ce gène, qui est apparenté aux protéines cdc20, p55, et fizzy, présente une plus faible homologie (44% d'identité) avec la protéine ccc52Ms, que les protéines fzf d'invertébrés et de levure ; ceci indique que la protéine de carotte appartient à un sous-groupe distinct de protéine à motifs répétés.

La recherche d'homologues de ccc52Ms dans une base de données du génome d'*Arabidopsis thaliana* a toutefois fait apparaître une séquence peptidique partielle déduite d'un clone génomique (AB005230), et présentant 64% d'identité avec ccc52Ms, ce qui montre l'existence d'homologues du gène ccc52Ms chez d'autres plantes.

La figure 1A représente l'alignement, effectué en utilisant le logiciel « PRETTYBOX », des séquences fzf de *Drosophila* (dmfzf), de *X. laevis* (xlfzf), hct1 de *Saccharomyces cerevisiae* (schct1), srw1 de *Schizosaccharomyces pombe* (srw1), et ccc52 de *Medicago sativa*, (ccc52ms) et d'*Arabidopsis thaliana* (ccc52at).

La figure 1B représente un dendrogramme de la famille des protéines à motifs WD40 répétés, qui montrent que les protéines ccc52 forment avec les autres protéines fzf, une sous-famille représentant une branche qui a évolué séparément de celles respectivement constituées par les protéines cdc20, p55, et fizzy.

La protéine ccc52Ms contient 7 domaines à motifs WD40 répétés, situés dans les portions centrale et C-terminale de la molécule (l'emplacement de ces domaines, numérotés de I à VII est indiqué sur la figure

1A au dessus de l'alignement des séquences). Ces domaines ne présentent que peu d'homologie entre eux, d'où l'on peut conclure qu'ils représentent des sites d'interaction avec des protéines différentes. Le dernier domaine (VII) comprend un site potentiel de liaison pour les cyclines.

Dans la partie N-terminale de la molécule, sont localisées une séquence peptidique (DRFIPSR) qui correspond à un motif conservé chez les protéines fzf ainsi que chez d'autres protéines à motifs WD40 répétés telles que cdc20, p55 et fizzy, ainsi qu'une séquence peptidique (AYTTLLRTALFG) qui correspond à un motif spécifique de la famille fzf (l'emplacement de ces motifs, respectivement indiqués A et B, sur la figure 1A au dessus de l'alignement des séquences).

3 sites potentiels de phosphorylation par des cdk (cyclin dependent kinases ou kinases cyclines-dépendantes), SPSR, SPVK et SPYK, sont localisés dans la portion N-terminale, respectivement aux positions 43-46, 128-131 et 144-147. Le premier et le second de ces sites sont également présents chez d'autres protéines fzf, tandis que le troisième paraît plus spécifique des protéines ccc52 de plantes.

Une séquence de 15 acides aminés RDNSPPPEPSPELR (encadrée sur la figure 1A), commençant au résidu 16, et correspondant à un motif de dégradation protéique PEST est également présente dans la portion N-terminale. Ce motif intervient probablement, par l'intermédiaire de la dégradation de ccc52, de réguler ses interactions avec d'autres protéines.

La structure de la protéine ccc52Ms est schématisée sur la figure 2.

La séquence de l'ADNc cloné par les inventeurs est représentée dans la liste de séquences en annexe sous le numéro SEQ ID NO:1 ; la séquence de la protéine ccc52Ms correspondante est représentée sous le numéro SEQ ID NO:2.

La région 3' non traduite du transcrit de cet ADN comprend 2 séquences AUUUA, qui correspondent à des séquences d'instabilité des ARNm, et peuvent donc jouer un rôle pour réguler la quantité de transcrits de ccc52.

5 Les Inventeurs ont étudié *in vivo* l'activité de la protéine ccc52Ms et ont montré qu'elle intervenait dans la régulation de la différenciation cellulaire, en favorisant l'endoréplication. En particulier, l'expression de la protéine ccc52Ms dans des plantes  
10 transgéniques induit chez celles-ci une augmentation de l'endoréplication et du niveau de ploïdie des cellules des plantes.

La présente invention a pour objet une protéine végétale à motifs WD40 répétés, dénommée ccc52,  
15 caractérisée en ce qu'elle appartient à la sous-famille fzs.

Selon un mode de réalisation préféré de la présente invention ladite protéine végétale présente au moins 45%, et de préférence au moins 55% d'identité avec  
20 le polypeptide de séquence SEQ ID NO:2 ou au moins 60% et de préférence au moins 70 % de similarité avec le polypeptide de séquence SEQ ID NO:2.

La présente invention englobe en particulier la protéine ccc52Ms, ses isoformes, ainsi que les  
25 protéines autologues de *Medicago* et les protéines orthologues d'autres végétaux, pouvant être rattachées à la famille des protéines fzs.

L'invention englobe également des protéines dérivées des protéines ccc52, par addition, délétion ou  
30 substitution d'un ou plusieurs acides aminés ou d'une ou plusieurs séquences d'acides aminés ; il peut s'agir par exemple de protéines dans lesquelles des modifications ont été apportées en dehors des régions fonctionnelles, ou bien de protéines dans lesquelles des modifications  
35 ont été apportées pour modifier leur activité, par

exemple de protéines stabilisées par délétion du motif PEST.

La présente invention a également pour objet un fragment d'acide nucléique purifié, caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une protéine ccc52, telle que définie ci-dessus, ou de sa séquence complémentaire. Dans ce cadre, la présente invention englobe en particulier les ADNc et les ADN génomiques des protéines ccc52.

Des fragments d'acide nucléiques, conformes à l'invention peuvent être aisément identifiés et clonés en criblant des banques d'ADNc ou d'ADN génomique de plantes à l'aide d'oligonucléotides dérivés de la séquence de ccc52Ms, et notamment d'oligonucléotides dérivés des régions de cette séquence spécifiques des protéines fzf, et en particulier des protéines ccc52.

Les protéines ccc52 peuvent être produites, en particulier, en exprimant ces séquences d'acide nucléique dans des cellules hôtes.

La présente invention a également pour objet l'utilisation d'une protéine ccc52, telle que définie ci-dessus ou d'une séquence d'acide nucléique codant pour tout ou partie de ladite protéine, ou de sa séquence complémentaire, pour réguler la différenciation et la prolifération de cellules végétales.

La présente invention a également pour objet l'utilisation d'une protéine de la sous-famille fzf ou d'une séquence d'acide nucléique codant pour tout ou partie de ladite protéine, ou de sa séquence complémentaire, pour réguler la différenciation et la prolifération de cellules végétales.

On peut citer parmi de telles protéines, la protéine fzf de drosophile ou la protéine fzf de levure.

La modification de l'expression et/ou de l'activité de protéines ccc52 dans des cellules de plantes permet de modifier le cycle cellulaire, en

favorisant soit la prolifération soit la différenciation, et de contrôler ainsi le processus de développement, afin d'obtenir par exemple une stimulation de l'embryogénèse somatique, d'augmenter la régénération *in vitro* de  
5 plantes à partir des cals, en augmentant la conversion en embryons, ou de favoriser le développement de certains organes, par exemple d'augmenter la productivité des tissus de réserve en augmentant leur endoploïdie.

On peut en particulier utiliser les séquences  
10 d'ADNc de protéines ccc52 ou des portions de ces séquences d'ADNc, ou de leur transcrits sens ou antisens ; il peut s'agir par exemple de la totalité d'une séquence codant pour une protéine ccc52Ms, ou d'une portion de cette séquence codante, et/ou de tout ou  
15 partie des régions 5' et 3' non traduites. Ces séquences peuvent être utilisées en orientation sens, ou si l'on souhaite inhiber l'expression de la protéine ccc52Ms dans une plante ou dans un tissu ou organe de celle-ci, en orientation antisens.

20 La présente invention englobe également des constructions d'ADN recombinant, contenant au moins une séquence d'acide nucléique conforme à l'invention.

Généralement, ladite séquence d'acide  
nucléique sera placée sous contrôle transcriptionnel d'un  
25 promoteur approprié.

Avantageusement, on pourra ainsi utiliser un promoteur fort, pour augmenter, dans les cellules hôte, les niveaux d'expression de la protéine ccc52 ; il pourra s'agir d'un promoteur inductible ou bien d'un promoteur  
30 constitutif ; il peut s'agir d'un promoteur ubiquitaire, ou d'un promoteur tissu-spécifique.

L'utilisation de promoteurs inductibles permet d'obtenir un blocage de la mitose, et l'induction de l'endoréplication au moment souhaité. L'utilisation de  
35 promoteurs tissu-spécifiques permet de cibler l'action de

la protéine ccc52 sur certains tissus et organes (par exemple, des tissus de réserve).

A titre d'exemples de promoteurs forts utilisables dans le cadre de la présente invention, on  
 5 citera : le promoteur CaMV35S [BENFLY et al., Science, 250, pp. 959-966, (1990)] ; le promoteur 35S ; les promoteurs *Agrobacterium tumefaciens* T-DNA : nopaline synthase, octopine synthase, mannopine synthase, 1', 2' [SANDERS et al., Nucleic Acid Res., 15, pp. 1543-1558,  
 10 (1987) ; HOOYKAAS and SCHILPEROORT, Plant. Mol. Biol., 19, pp. 15-38, (1992)].

A titre d'exemples de promoteurs inductibles utilisables dans le cadre de la présente invention, on  
 15 citera : le promoteur inductible par la tétracycline [WEINMANN et al., Plant J., 5, pp. 559-569, (1994)] ; le promoteur inductible par le cuivre [METT et al., Transgenic Res., 5, pp. 105-113, (1996)] ; le promoteur inductible par les glucocorticoïdes [AOYAMA et CHUA, Plant. J., 11, pp. 605-612, (1997)].

20 A titre d'exemples de promoteurs tissu-spécifiques utilisables dans le cadre de la présente invention, on citera : le promoteur endosperme-spécifique [OPSAHL-FERSTAD et al., Plant J., 12, pp. 235-246, (1997) ; DOAN et al., Plant Mol. Biol., 31, pp. 877-886,  
 25 (1996) ; les promoteurs nodosités-spécifiques (*enod12A/B* ou leghémoglobine) [TRINH et al., Plant Cell Reports, (17, pp. 345-355, (1998) ; VIJN et al., Plant Mol. Biol., 28, pp. 1103-1110, (1995)] ou bien des promoteurs précoces inductibles par le facteur Nod et des promoteurs  
 30 tardifs (promoteur de la cycline D ou des nodulines tardives (type leghémoglobine) et promoteurs régulés par des hormones, tels que *parA/B* [TAKAHASHI et al., Proc. Natl. Acad. Sci. USA, 87, pp. 8013-8016, (1990)], *GH3* [LIU et al., Plant Cell, 6, pp. 645-657, (1994)].

35 L'Invention englobe en particulier des vecteurs recombinants portant au moins un insert

contenant un fragment d'ADN conforme à l'invention. Ces vecteurs sont utilisables pour transformer des cellules hôtes.

5 L'Invention a également pour objet des cellules et des organismes pluricellulaires transformés par au moins une séquence d'acide nucléique conforme à l'invention ; il s'agit en particulier de cellules végétales ou de végétaux.

10 La présente invention sera mieux comprise à l'aide du complément de description qui va suivre, et se réfère à des exemples non limitatifs illustrant l'identification, le clonage et l'expression du gène ccc52Ms.

15 EXEMPLE 1 : CLONAGE ET SEQUENCAGE D'UN ADNC de ccc52Ms.

Un clone d'ADNC de ccc52Ms a été obtenu par criblage différentiel à partir d'une banque d'ADNC de nodosités de *Medicago sativa*, fortement stimulées pendant l'organogénèse nodulaire.

20 Le protocole suivant a été utilisé :

L'ADNC de ccc52Ms de *M. sativa* est isolé par la technique de DD-RT-PCR (Differential Display RT-PCR) [LIANG et PARDEE, Science, 257, pp. 967-971, (1992)], en utilisant les kits RNAimage (GenHunter Corporation). Les  
25 échantillons d'ARN sont isolés à partir de la zone racinaire sensible au facteur Nod de jeunes plants de *M. sativa* (croissance dans un milieu limité en nitrate), en l'absence de bactéries ou inoculés par des souches de *R. meliloti* Nod+ (EK1433) ou Nod- (EK133) pendant  
30 4 jours. Le fragment de DD-RT-PCR ccc52Ms, présentant une augmentation de l'expression des nodosités, est cloné dans le vecteur de clonage pCT-TRAP (GenHunter Corporation) et utilisé comme sonde pour l'isolement des clones complets à partir d'une banque d'ADNC de nodosités  
35 de *M. sativa* sp. *varia* A2, construite dans  $\lambda$ -ZAP



(Stratagène) (CRESPI et al., EMBO J., 1994, 13, 5099-5112).

Sept clones d'ADNc, obtenus à partir de  $2.10^5$  phages et coupés *in vivo*, conformément au protocole de Stratagène, représentent 2 types d'ADNc différant l'un de l'autre uniquement au niveau de 4 acides aminés (16R-G, 17D-N, 33S-N, 52R-G) et de la longueur du fragment 3'UTR. Une identité de 99% des clones, au niveau de la séquence en acides aminés, suggère qu'ils représentent des allèles du même gène chez *M. sativa* tétraploïde allogame.

Le séquençage des ADNc de *ccc52Ms* est effectué avec le système ABIprism de Perkin-Elmer, utilisant les amorces T3 et T7 du vecteur Bluescript (Stratagène) et les amorces spécifiques suivantes :

15 52'L : GAAATCCACCACCGACTT,  
P55A2-L : AGCAACCATTACACCTCACC  
p55B-L : TTTGGGGGTTGATGATTGTG  
p55CL : CTCTCTACCGTTCTATCTCTTGGA  
P55B-R : CCTCTGATTTGTGTCCTGA  
20 P55-R : AGCTTTGACATGGTGGGGTATCTCC  
P55CR : GGTAAGATGCTACTTTGGTGGTGT.

Les clones génomiques *ccc52Ms* et *ccc52Mt* sont isolés à partir de banques génomiques de *M. sativa* cv. Nagyszénasi et *M. truncatula* ecotype GHOR, en utilisant l'ADNc de *ccc52Ms* comme sonde d'hybridation. Ces banques génomiques sont construites par digestion partielle de l'ADN génomique avec l'enzyme de restriction MboI et le clonage des fragments d'ADN de taille comprise entre 15 et 20 Kb dans le site BamHI de  $\lambda$ -EMBL4.

30 EXEMPLE 2 : IDENTIFICATION DE LA FAMILLE DU GENE *ccc52Ms* DANS *Medicago* ET SON EXPRESSION DANS DIFFERENTS ORGANES VEGETAUX.

L'existence de copies multiples du gène *ccc52* est recherchée par hybridation de type Southern dans des cultivars tétraploïdes de *M. sativa* Nagyszénasi et Cardinal et chez *M. truncatula* diploïde autogame, une

plante modèle dans la recherche sur les légumes (Fig. 3A).

L'ADN de plante est isolé à partir des feuilles jeunes, en utilisant le kit d'extraction Nucleon  
5 PhytoPure DNA (Amersham).

Les échantillons d'ADN sont digérés par EcoRI et transférés sur membrane de nylon Biotrans (+) (ICN).

L'hybridation Southern est réalisée conformément aux protocoles publiés, dans des conditions  
10 stringentes à 65°C (hybridation dans un tampon CG ; lavage : 2 x SSC, 0,1% SDS pendant 2 fois 15 mn, puis 0,5 x SSC, 0,1% SDS pendant 2 fois 30 min.) (SAMBROOK, Molecular Cloning: A Laboratory Manual 2nd edn., Cold Spring Harbor Laboratory Press, New York, (1989) ;  
15 AUSUBEL, Current Protocols in Molecular Biology, (1989)].

L'expression de *ccc52Ms* est étudiée par analyse Northern.

L'ARN total est isolé à partir de différents organes de *M. sativa* cultivar Sitel :

- 20 - à partir des racines, inoculées pendant 4 jours avec le mutant Nod- *R. meliloti* (EK133) et avec la souche surproductrice de facteurs Nod (EK1433) ;
- à partir des nodosités, 12, 19, 23 et 30 jours après une infection par *R. meliloti*, et,
- 25 - à partir des tiges, des hypocotyles, des feuilles, des bourgeons, des fleurs, des racines de plants de 3 jours, de 7 jours, des racines privées d'azote et ne présentant pas de pointes racinaires, des racines de 7 jours, sans pointes racinaires, mises en
- 30 culture en présence de nitrate, des nodosités spontanées développées en l'absence de *R. meliloti* et des pointes racinaires ou de culture de cellules de *M. sativa* sp. *varia* A2.

100 mg de plante, collectés sous azote  
35 liquide, sont utilisés pour l'extraction de l'ARN (RNeasy Plant, Quiagen).

10 µg d'ARN est chargé par ligne sur un gel dénaturant (formaldéhyde) [SAMBROOK, Molecular Cloning: A Laboratory Manual 2nd edn., Cold Spring Harbor Laboratory Press, New York, (1989)].

5 Le transfert d'ADN est effectué dans une solution de transfert 10 x SSC [CHOMCZYNSKI et al., Analytical Biochemistry, 221, pp. 303-305, (1994)].

Aussi bien dans le cas de l'hybridation Southern que dans le cas de l'hybridation Northern, le  
10 fragment d'ADNc *ccc52Ms* est marqué avec du [ $\alpha$ -<sup>32</sup>P] dCTP (Kit Megaprim, Amersham). L'hybridation avec la sonde *Msc27* sert de contrôle pour le chargement de l'ARN [SAVOURE et al., EMBO J., 13, pp. 1093-1102, (1994)].

a) Les résultats sont illustrés par la figure  
15 3A, qui représente l'hybridation avec la sonde de l'ADN génomique de *M. sativa* ou *M. truncatula* digéré avec EcoRI (E) ; on observe que la sonde s'hybride avec différents fragments EcoRI, ce qui indique que *ccc52Ms* représente chez *Medicago*, une petite famille  
20 multigénique.

b) La figure 3B illustre les résultats obtenus avec l'ARN total de racines inoculées avec le mutant Nod<sup>-</sup> EK133 de *R. meliloti* (piste 1), ou avec la souche EK1433 surproductrice de facteurs Nod (piste 2) et avec l'ARN  
25 extrait des nodosités, 12, 19, 23 et 30 jours après infection avec *R. meliloti* (pistes 3-6, respectivement). La sonde *Msc27* a été utilisée comme contrôle de la présence d'ARN.

Ces résultats montrent qu'on n'observe dans  
30 l'ARN total de racines, qu'une faible quantité de transcrits, ce qui reflète la faible proportion des cellules impliquées dans l'organogénèse des nodosités par rapport au nombre total de cellules des racines. En revanche, dans les nodosités de différents âges on  
35 observe un niveau élevé de transcription, qui reflète la

persistance des méristèmes apicaux et des zones de différenciation.

c) la figure 3C montre les résultats obtenus avec les ARN totaux de : 1: culture de cellules de *M. sativa* sp. *varia* A2, 2: tiges, 3: hypocotyles, 4: feuilles, 5: bourgeons floraux, 6: fleurs, 7: racines de pousses de 3 jours, 8: racines de pousses de 7 jours, privées d'azote, dépourvues d'extrémité, 9: pointes racinaires de 7 jours, cultivées en présence de nitrates, 10: dépourvues d'extrémité, 10: nodosités spontanées développées en absence de *R. melioli*, 11: nodosités fixatrices d'azote, 12: extrémités de pointes racinaires.

Cette analyse montre que l'expression du *ccc52Ms* n'est pas limitée aux nodosités, bien que cet organe soit celui qui contienne le niveau de transcrits le plus élevé. Ces transcrits sont présents en quantités variables pratiquement dans tous les organes, ce qui indique que cette protéine intervient dans le développement de chacun d'entre eux. Mis à part les nodosités, le niveau de transcription est également élevé chez les jeunes pousses, et dans les cultures cellulaires, où l'on détecte en outre un ARNm de plus petite taille qui peut correspondre soit à une polyadénylation différente, soit à l'expression d'une copie homologue du gène.

Des analyses par hybridation *in situ* ont également été effectuées, et montrent que l'ARNm de *ccc52Ms* est localisé principalement dans la zone de différenciation, et en particulier à l'interface entre les zones II et III de la nodosité, qui sont les régions où la différenciation est la plus active.

Les hybridations radioactives sont réalisées selon la technique de SAVOURE et al. [EMBO J., 13, pp. 1093-1102, (1994)] alors que les hybridations non-radioactives (marquage à la digoxigénine) sont réalisés

conformément aux protocoles établis par le fabricant (Boehringer-Mannheim).

Parallèlement, on observe dans les mêmes zones une expression des cyclines de type G1 et mitotiques, ainsi que de l'histone H3 spécifique de la phase S.

Ceci indique que ccc52Ms intervient dans la régulation du cycle cellulaire, probablement d'une manière similaire à ses homologues de la levure et de la drosophile, c'est-à-dire par l'intermédiaire de la protéolyse de cyclines mitotiques, qui inhibe la mitose et induit des cycles d'endoreduplication.

EXEMPLE 3 : EXPRESSION DE ccc52Ms CHEZ *Schizosaccharomyces pombe*

L'expression de ccc52Ms a été étudiée chez *S. pombe* chez qui un homologue fonctionnel *srw1* a été décrit récemment (YAMAGUCHI, publication précitée). Le gène codant pour ccc52Ms a été cloné dans le plasmide pREP1 sous contrôle du promoteur *Pnmt1* répressible par la thiamine.

L'ADNc de ccc52Ms obtenu après coupure de  $\lambda$ -ZAP (Stratagène) est digéré avec AgeI et partiellement avec EcoRV. Le fragment AgeI-EcoRV de 1,6 kb représentant la région codante, à l'exception des 4 premiers codons, est cloné dans un vecteur SKII Bluescript (Stratagène) digéré par XmaI (compatible avec AgeI) et EcoRV. A partir de ce plasmide (pSK52B), l'ADNc de ccc52Ms est coupé par digestion BamHI-EcoRV et cloné dans les sites BamHI-SmaI du plasmide pREP1 [MAUNDRELL et al., Gene, 123, pp. 127-30, (1993)]. Pour générer un cadre de lecture en phase avec le codon de traduction ATG présent dans le vecteur, derrière le promoteur *nmtI*, l'ADN est digéré avec BamHI et l'extrémité 5' est complétée en présence d'enzyme de Klenow et de dNTPs. La religature des extrémités à bouts francs entraîne une fusion correcte, également vérifiée par séquençage. Ce plasmide, dénommé pREP52, est transformé dans des cellules compétentes

*S. pombe* SP-Q01 et les transformants sont sélectionnés sur des plaques d'agar EMM-thiamine, en utilisant le kit ESP<sup>®</sup> (Stratagène). Les vecteurs pREP1 [MAUNDRELL et al., Gene, 123, pp. 127-30, (1993)] et pESP1 (Stratagène) sont  
 5 utilisés comme contrôles négatifs ; un contrôle positif est constitué de *srw1* cloné dans pREP1 [YAMAGUSHI et al, Mol. Biol. Cell., 8, pp. 2475-2486, (1997)], transformés dans *S. pombe* SP-Q01.

Les transformants purifiés *S. pombe* SP-Q01  
 10 sont cultivés dans 2 ml de milieu EMM-thiamine 5  $\mu$ M pendant 32 h à 30°C. Les cellules sont lavées 2 fois avec 10 ml d'eau stérile et remises en suspension dans 5 ml de milieu EMM. Les suspensions cellulaires sont divisées en deux moitiés : 2,5 ml sont cultivées avec de la thiamine  
 15 et 2,5 ml sont cultivées sans thiamine, à 30°C. Des aliquots de cultures sont prélevés après 16 h et 24 h de culture et fixés avec de l'éthanol, colorés avec du DAPI ou de l'iodure de propidium pour une analyse en cytométrie de flux et en microscopie [BEACH et al, Curr.  
 20 Genet., 10, pp. 297-311, (1985)].

En présence de thiamine, l'expression de *ccc52Ms* est réprimée et on observe une croissance normale (Figure 4A).

En l'absence de thiamine (Figure 4B), la  
 25 surexpression de *ccc52Ms* entraîne l'inhibition de la croissance de *S. pombe*, qui s'accompagne d'une endoréplication comme le montre la figure 4D, qui montre la présence de noyaux à 4C, qui n'est pas observée dans les cellules contrôles *S. pombe*, portant le vecteur vide  
 30 pREP1 (figure 4C).

La morphologie des cellules est également modifiée en l'absence de thiamine (figure 4F). On observe un allongement des cellules et une augmentation de la taille des noyaux, identiques à ceux observés lors de la  
 35 surexpression de l'homologue *srw1* [YAMAGUSHI et al, Mol. Biol. Cell., 8, pp. 2475-2486, (1997)], tandis qu'aucun

changement morphologique n'est observé lorsque *S. pombe* ne porte que le vecteur pREP1 (figure 4E).

#### EXEMPLE 4 : OBTENTION DE PLANTES TRANSGENIQUES TRANSFORMEES PAR LE GENE *ccc52Ms*.

5                   1° Expression d'un transcrit antisens et son action sur le niveau de ploïdie de *Medicago truncatula*.

Dans un premier temps, le niveau de ploïdie de différents organes de *Medicago truncatula* (plante naturellement diploïde) a été déterminé, par cytométrie  
10 de flux. Les résultats sont illustrés par la figure 5A, au niveau des pétioles.

La technique utilisée est la suivante :

l'ADN nucléaire de plantes fraîchement récoltées est analysé par cytométrie de flux (EPICS V,  
15 Coulter), conformément à la méthode de Brown et al., (A laboratory guide for Cellular and Molecular plant Biology, 1991, 326-345, ed. Negrutiu et al., Birkhäuser, Basel), modifiée de telle sorte que les noyaux soient colorés avec du DAPI à une concentration finale de  
20 5 µg/ml. Le tampon nucléaire I est utilisé à 1% de Triton X-100 pour les nodosités.

Dans les jeunes pousses, on trouve une quantité d'ADN de 2C à 8C dans la racine et le cotylédon, alors que l'hypocotyle contient également des noyaux à  
25 16C. Chez les plantes adultes poussées en serre, les feuilles sont diploïdes, contenant 95% de noyaux à 2C et 5% de noyaux à 4C. Dans les pétioles et les nodosités, des noyaux de 2C à 32C ont été détectés. Toutefois, le pétiole contient majoritairement des noyaux à 2C, alors  
30 que les nodosités contiennent majoritairement des noyaux à 4C.

Un fragment de 1,2kb contenant les ¾ de la séquence codante de *ccc52Ms*, a été placé en orientation antisens sous contrôle du promoteur 35S, dans un vecteur  
35 binaire obtenu à partir du vecteur pGPTV-BAR, portant le gène *bar* comme marqueur de sélection et des sites de

clonage multiple. Cette construction est obtenue en insérant le promoteur 35S dans un fragment HindIII-XbaI (obtenu à partir de pBI121, Clontech), dans les sites HindIII-XbaI du vecteur pGPTV-BAR codant le gène de  
 5 résistance à l'herbicide BASTA. Le gène uidA est ensuite éliminé du plasmide pGPTV-BAR par digestion XbaI-SstI au niveau du site multiple de clonage.

Pour obtenir la construction antisens de ccc52Ms, un fragment SstI-PvuII de 1,2 kb est cloné au  
 10 niveau des sites SmaI-SstI du vecteur binaire précité.

Ces plasmides ainsi qu'un plasmide contrôle, contenant le gène gus ont été introduits dans *Agrobacterium tumefaciens* (EHA105) par électroporation et utilisés pour transformer *Medicago truncatula* R108-1  
 15 selon le protocole décrit par HOFFMANN et al. [Mol. Plant Microbe Interaction, 10, pp. 307-315, (1997)] ; TRINH ET AL. [Plant Cell Reports, 17, pp. 345-355, (1998)].

Le niveau de ploïdie des plantes transgéniques obtenues a été analysé, et le niveau de transcrits  
 20 endogènes a été analysé par RT-PCR.

Dans la mesure où le fragment SstI-PvuII de 1,2 kb utilisé pour la construction antisens ne contient pas le dernier quart de la séquence codante, les taux de transcrit endogène ccc52Mt peuvent être analysés dans les  
 25 lignées transgéniques par RT-PCR en utilisant la paire d'amorces P55CL et P55CR comme illustré aux figures 5E et 5F.

Sur 37 plantes transgéniques, 3 ont montré une endoploïdie réduite, et une autre (A4) une endoploïdie  
 30 significativement plus faible. C'est également dans cette plante que le niveau d'expression des transcrits endogènes de ccc52Ms est le plus faible, comme le montre la figure 5F. Dans cette plante, les pétioles contenaient 48% de noyaux 2C, 50,2% de noyaux 4C, 1,65% de noyaux 8C,  
 35 alors que les pétioles des plantes témoin contiennent 51,5% de noyaux 2C, 31,8% de noyaux 4C, 14,1% de



noyaux 8C, 2,6% de noyau 16C, et 0,178% de noyaux 32C. En outre, cette plante produit une quantité de graines significativement inférieure à celle des plantes témoin.

Le fait qu'une réduction de l'endoploidie n'ait jamais été observée auparavant chez d'autres plantes transgéniques, permet d'attribuer ce phénomène à l'altération de l'expression de *ccc52Ms*, et non à un effet secondaire de la transgénèse.

2° Expression de la protéine *ccc52Ms* dans des plantes transgéniques.

Des vecteurs d'expression contenant le gène *ccc52Ms* sous contrôle du promoteur 35S, ainsi que des vecteurs d'expression comprenant le gène *ccc52Ms*, sous contrôle d'un promoteur tissu-spécifique, ont été construits selon le protocole suivant :

Pour l'expression tissu-spécifique de *ccc52Ms*, l'ADNc est placé sous le contrôle des promoteurs *enod12AMs* et *Srg1b3* décrits dans TRINH et al. [Plant Cell Reports, 17, pp. 345-355, (1998)], sauf que pISV-BMCS, un dérivé de pISV2301, et seulement un fragment du promoteur *enod12AMs* de 0,3 kb considéré comme suffisant pour une expression nodosité-spécifique [VIJN et al., Plant Mol. Biol., 28, pp. 1103-1110, (1995)], sont utilisés.

Construction de pISV-BMCS : pISV2301 est digéré par HindIII et SstI pour éliminer la séquence du promoteur 2X35S-AMV, qui est remplacé par l'oligo BMCS double-brin

AGCTTCCCGGGGAGCTCTAGACTCGAGCAGCT BMCS

AGGCCCTCGAGATCTGAGCTCG

SmaI SstI XbaI XhoI.

pISV-BMCS12A est construit par clonage d'un fragment du promoteur *enod12AMs* de 0,3 kb, généré par PCR, en utilisant le plasmide pPR89 (matrice) [BAUER et al., Plant J., 10, pp. 91-105, (1996)] et les amorces PE12A-L et PE12A-R, dans pISV-BMCS.

pISV-BMCS-LB3 est construit par digestion de pISV-BMCS avec HindIII-SstI et clonage d'un fragment HindIII-SstI contenant le promoteur leghémoglobine de *Sesbania rostrata* à partir de pLP32 [TRINH et al, Plant Cell Reports, 17, pp. 345-355, (1998)].

Ces vecteurs ont été utilisés pour transformer *Medicago truncatula* selon le protocole décrit ci-dessus pour les séquences antisens.

Lors de la régénération des plantes transgéniques, on observe une conversion des cals en embryons significativement plus importante chez les plantes transformées avec les constructions exprimant le gène ccc52Ms, que chez les plantes transformées avec la construction témoin, ce qui indique un effet positif de ccc52Ms sur l'embryogénèse somatique.

## LISTE DE SEQUENCES

## (1) INFORMATIONS GENERALES:

## (i) DEPOSANT:

- (A) NOM: CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE  
(CNRS)  
(B) RUE: 3 rue MICHEL-ANGE  
(C) VILLE: PARIS  
(E) PAYS: FRANCE  
(F) CODE POSTAL: 75794

(ii) TITRE DE L'INVENTION: GENES INTERVENANT DANS LA REGULATION DE LA DIVISION CELLULAIRE CHEZ LES VEGETAUX ET LEURS UTILISATIONS.

(iii) NOMBRE DE SEQUENCES: 2

## (iv) FORME DECHIFFRABLE PAR ORDINATEUR:

- (A) TYPE DE SUPPORT: Floppy disk  
(B) ORDINATEUR: IBM PC compatible  
(C) SYSTEME D' EXPLOITATION: PC-DOS/MS-DOS  
(D) LOGICIEL: PatentIn Release #1.0, Version #1.30 (OEB)

## (2) INFORMATIONS POUR LA SEQ ID NO: 1:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

- (A) LONGUEUR: 2006 paires de bases  
(B) TYPE: nucléotide  
(C) NOMBRE DE BRINS: simple  
(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: ADNc

## (ix) CARACTERISTIQUE:

- (A) NOM/CLE: CDS  
(B) EMBLEMMENT: 182..1609

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 1:

```

GATTCGGCAC GAGGAAGAAA CAAAGAAACT CTCTCTCTCT ATTTCTTTCT CTCTGCACAA      60
TTTTCGAGTA GTGTTATTTT TTAATAAAAA ATTAATTAAT TTTT TTTTAT ATAAAAGCCG      120
TGCAAAAAAT TCTTTTACAG CGTTCTTTTT TCCCCGGGAA AAAAATTAAC ACAGCTCCGC      180
C ATG GAC GGA ACC GGT AAT CGA AAT CCA CCA CCG ACT TCC ACC GTC      226
Met Asp Gly Thr Gly Asn Arg Asn Pro Pro Pro Thr Ser Thr Val
  1           5           10           15

AGA GAC AAT TCT CCA CCG CCT GAG CCA TCA CCG GAG AGT CTC CGT CAT      274
Arg Asp Asn Ser Pro Pro Pro Glu Pro Ser Pro Glu Ser Leu Arg His
          20           25           30

GTA AGC CGT ATG ATC AAC AGC AAC CAT TAC ACC TCA CCT TCT CGA ACA      322
Val Ser Arg Met Ile Asn Ser Asn His Tyr Thr Ser Pro Ser Arg Thr
          35           40           45

ATC TAC TCC GAT AGG TTC ATT CCG AGT AGA TCT GCT TCG AAA TTC GCT      370
Ile Tyr Ser Asp Arg Phe Ile Pro Ser Arg Ser Ala Ser Lys Phe Ala
          50           55           60

```

TTG TTT GAT ATC AAT ACT CCG ACA GAA GGA CGC GAT GAT AGT TCC AGC Leu Phe Asp Ile Asn Thr Pro Thr Glu Gly Arg Asp Ser Ser Ser 65 70 75	418
GCT TAT ACG ACT CTT CTG AGA ACG GCG TTG TTT GGA CCG GAT GTT GCC Ala Tyr Thr Thr Leu Leu Arg Thr Ala Leu Phe Gly Pro Asp Val Ala 80 85 90 95	466
GGT CCG GTT ACG CCG GAA AAA ACC GAC TCG CCG TCG ATG ACA TTG CCG Gly Pro Val Thr Pro Glu Lys Thr Asp Ser Pro Ser Met Thr Leu Pro 100 105 110	514
AAT AGG AAT ATT TTT AGG TAT AAG ACG GAG ACG AGA CAG TCC ATG CAC Asn Arg Asn Ile Phe Arg Tyr Lys Thr Glu Thr Arg Gln Ser Met His 115 120 125	562
TCG CTT TCG CCG TTT ATG GAT GAT GAT TTT GTT CCT GGT GTT AAT CAT Ser Leu Ser Pro Phe Met Asp Asp Asp Phe Val Pro Gly Val Asn His 130 135 140	610
AGT CCG GTT AAG GCT CCT AGG AAG GTT CCT CGA TCG CCT TAT AAG GTT Ser Pro Val Lys Ala Pro Arg Lys Val Pro Arg Ser Pro Tyr Lys Val 145 150 155	658
TTG GAT GCA CCT GCT TTG CAA GAT GAT TTT TAT CTG AAT CTG GTA GAT Leu Asp Ala Pro Ala Leu Gln Asp Asp Phe Tyr Leu Asn Leu Val Asp 160 165 170 175	706
TGG TCT TCA CAC AAT GTG TTG GCT GTT GGT TTG GGT AAC TGT GTC TAT Trp Ser Ser His Asn Val Leu Ala Val Gly Leu Gly Asn Cys Val Tyr 180 185 190	754
CTC TGG AAT GCT TGT AGC AGC AAG GTA ACT AAA TTA TGT GAT TTG GCG Leu Trp Asn Ala Cys Ser Ser Lys Val Thr Lys Leu Cys Asp Leu Gly 195 200 205	802
GTT GAT GAT TGT GTT TGT TCT GTT GGT TGG GCT CAA CGT GGT ACT CAT Val Asp Asp Cys Val Cys Ser Val Gly Trp Ala Gln Arg Gly Thr His 210 215 220	850
CTT GCT GTT GGA ACT AAC AAT GGT AAA GTT CAG ATT TGG GAT GCA GCA Leu Ala Val Gly Thr Asn Asn Gly Lys Val Gln Ile Trp Asp Ala Ala 225 230 235	898
AGA TGC AAG AAG ATA AGA TCA ATG GAG GGC CAT CGG TTA CGT GTC GGG Arg Cys Lys Lys Ile Arg Ser Met Glu Gly His Arg Leu Arg Val Gly 240 245 250 255	946
GCC TTG GCC TGG AGT TCA TCT CTT TTG TCT TCT GGT GGA CGG GAT AAG Ala Leu Ala Trp Ser Ser Ser Leu Leu Ser Ser Gly Gly Arg Asp Lys 260 265 270	994
AAT ATT TAT CAA CGA GAT ATA CGC ACA CAA GAA GAT TTT GTT AGT AAA Asn Ile Tyr Gln Arg Asp Ile Arg Thr Gln Glu Asp Phe Val Ser Lys 275 280 285	1042
CTG TCA GGA CAC AAA TCA GAG GTT TGT GGA CTG AAG TGG TCA TAT GAT Leu Ser Gly His Lys Ser Glu Val Cys Gly Leu Lys Trp Ser Tyr Asp 290 295 300	1090
AAC CGT GAG TTG GCA TCT GGA GGA AAT GAC AAC AAA TTG TTT GTT TGG Asn Arg Glu Leu Ala Ser Gly Gly Asn Asp Asn Lys Leu Phe Val Trp 305 310 315	1138

AAT CAA CAC TCA ACC CAG CCT GTC CTC AAG TAC TGT GAG CAC ACA GCA Asn Gln His Ser Thr Gln Pro Val Leu Lys Tyr Cys Glu His Thr Ala 320 325 330 335	1186
GCT GTT AAA GCT ATT GCA TGG TCT CCT CAT CTT CAT GGA CTT CTT GCA Ala Val Lys Ala Ile Ala Trp Ser Pro His Leu His Gly Leu Leu Ala 340 345 350	1234
TCT GGA GGA GGA ACT GCA GAT AGA TGT ATT CGT TTT TGG AAT ACA ACC Ser Gly Gly Gly Thr Ala Asp Arg Cys Ile Arg Phe Trp Asn Thr Thr 355 360 365	1282
ACA AAC TCA CAC CTT AGC TGT ATG GAC ACT GGA AGT CAG GTT TGC AAT Thr Asn Ser His Leu Ser Cys Met Asp Thr Gly Ser Gln Val Cys Asn 370 375 380	1330
CTT GTC TGG TCC AAA AAT GTC AAC GAA CTA GTA AGC ACA CAT GGG TAC Leu Val Trp Ser Lys Asn Val Asn Glu Leu Val Ser Thr His Gly Tyr 385 390 395	1378
TCC CAG AAC CAG ATT ATT GTT TGG AGA TAC CCC ACT ATG TCA AAG CTG Ser Gln Asn Gln Ile Ile Val Trp Arg Tyr Pro Thr Met Ser Lys Leu 400 405 410 415	1426
GCG ACT CTT ACC GGC CAT ACT TAT AGG GTT CTC TAT CTT GCC ATC TCT Ala Thr Leu Thr Gly His Thr Tyr Arg Val Leu Tyr Leu Ala Ile Ser 420 425 430	1474
CCA GAT GGA CAG ACT ATT GTA ACT GGA GCT GGA GAT GAA ACG CTT AGG Pro Asp Gly Gln Thr Ile Val Thr Gly Ala Gly Asp Glu Thr Leu Arg 435 440 445	1522
TTC TGG AAT GTT TTC CCT TCC CCT AAA TCA CAG AAT ACT GAA AGT GAA Phe Trp Asn Val Phe Pro Ser Pro Lys Ser Gln Asn Thr Glu Ser Glu 450 455 460	1570
ATC GGA GCA TTA TCT CTT GGA AGA ACT ACT ATC AGG TGA TTGATCCTGG Ile Gly Ala Leu Ser Leu Gly Arg Thr Thr Ile Arg * 465 470 475	1619
CGTTGCAGCC CAATCATGTG GCATATTTCT AAGTTTGGGT TGCTGTGTAG AACTAAATTT	1679
CTGAGCGGAG AACACCATGG TGGAAAAACC TTGAATATAA AAACACCACC AAAGTAGCAT	1739
CTTTACCAAC TGGGAGAGCC TTGGAGGGAG CTATAAAAGT TTTGATATGG CTGCCGGTGA	1799
TATTCCTGCA TTCATGTGTA GTCTCATTTT ATATTGAAAA GATGATAACA AATGGGTAAT	1859
TTATTGTCTT GGA CT TATAC ATGCATTGAT GGAGTTGTAG CCAAGTTTTT TTATTACTCT	1919
TTTTTTCTTT CTTCTTTTTG ATAGTGCTCT CCTGCATTAT TTATATAATT TTAAGATGCG	1979
TTAACAGAGA AAAAAAAAAA AAAAAAA	2006

## (2) INFORMATIONS POUR LA SEQ ID NO: 2:

## (i) CARACTERISTIQUES DE LA SEQUENCE:

(A) LONGUEUR: 476 acides aminés

(B) TYPE: acide aminé

(D) CONFIGURATION: linéaire

(ii) TYPE DE MOLECULE: protéine

## (xi) DESCRIPTION DE LA SEQUENCE: SEQ ID NO: 2:

Met Asp Gly Thr Gly Asn Arg Asn Pro Pro Pro Thr Ser Thr Val Arg  
 1 5 10 15  
 Asp Asn Ser Pro Pro Pro Glu Pro Ser Pro Glu Ser Leu Arg His Val  
 20 25 30  
 Ser Arg Met Ile Asn Ser Asn His Tyr Thr Ser Pro Ser Arg Thr Ile  
 35 40 45  
 Tyr Ser Asp Arg Phe Ile Pro Ser Arg Ser Ala Ser Lys Phe Ala Leu  
 50 55 60  
 Phe Asp Ile Asn Thr Pro Thr Glu Gly Arg Asp Asp Ser Ser Ser Ala  
 65 70 75 80  
 Tyr Thr Thr Leu Leu Arg Thr Ala Leu Phe Gly Pro Asp Val Ala Gly  
 85 90 95  
 Pro Val Thr Pro Glu Lys Thr Asp Ser Pro Ser Met Thr Leu Pro Asn  
 100 105 110  
 Arg Asn Ile Phe Arg Tyr Lys Thr Glu Thr Arg Gln Ser Met His Ser  
 115 120 125  
 Leu Ser Pro Phe Met Asp Asp Asp Phe Val Pro Gly Val Asn His Ser  
 130 135 140  
 Pro Val Lys Ala Pro Arg Lys Val Pro Arg Ser Pro Tyr Lys Val Leu  
 145 150 155 160  
 Asp Ala Pro Ala Leu Gln Asp Asp Phe Tyr Leu Asn Leu Val Asp Trp  
 165 170 175  
 Ser Ser His Asn Val Leu Ala Val Gly Leu Gly Asn Cys Val Tyr Leu  
 180 185 190  
 Trp Asn Ala Cys Ser Ser Lys Val Thr Lys Leu Cys Asp Leu Gly Val  
 195 200 205  
 Asp Asp Cys Val Cys Ser Val Gly Trp Ala Gln Arg Gly Thr His Leu  
 210 215 220  
 Ala Val Gly Thr Asn Asn Gly Lys Val Gln Ile Trp Asp Ala Ala Arg  
 225 230 235 240  
 Cys Lys Lys Ile Arg Ser Met Glu Gly His Arg Leu Arg Val Gly Ala  
 245 250 255  
 Leu Ala Trp Ser Ser Ser Leu Leu Ser Ser Gly Gly Arg Asp Lys Asn  
 260 265 270  
 Ile Tyr Gln Arg Asp Ile Arg Thr Gln Glu Asp Phe Val Ser Lys Leu  
 275 280 285  
 Ser Gly His Lys Ser Glu Val Cys Gly Leu Lys Trp Ser Tyr Asp Asn  
 290 295 300  
 Arg Glu Leu Ala Ser Gly Gly Asn Asp Asn Lys Leu Phe Val Trp Asn  
 305 310 315 320  
 Gln His Ser Thr Gln Pro Val Leu Lys Tyr Cys Glu His Thr Ala Ala  
 325 330 335

Val Lys Ala Ile Ala Trp Ser Pro His Leu His Gly Leu Leu Ala Ser  
 340 345 350  
 Gly Gly Gly Thr Ala Asp Arg Cys Ile Arg Phe Trp Asn Thr Thr Thr  
 355 360 365  
 Asn Ser His Leu Ser Cys Met Asp Thr Gly Ser Gln Val Cys Asn Leu  
 370 375 380  
 Val Trp Ser Lys Asn Val Asn Glu Leu Val Ser Thr His Gly Tyr Ser  
 385 390 395 400  
 Gln Asn Gln Ile Ile Val Trp Arg Tyr Pro Thr Met Ser Lys Leu Ala  
 405 410 415  
 Thr Leu Thr Gly His Thr Tyr Arg Val Leu Tyr Leu Ala Ile Ser Pro  
 420 425 430  
 Asp Gly Gln Thr Ile Val Thr Gly Ala Gly Asp Glu Thr Leu Arg Phe  
 435 440 445  
 Trp Asn Val Phe Pro Ser Pro Lys Ser Gln Asn Thr Glu Ser Glu Ile  
 450 455 460  
 Gly Ala Leu Ser Leu Gly Arg Thr Thr Ile Arg \*  
 465 470 475

## REVENDICATIONS

1) Protéine végétale à motifs WD40 répétés, caractérisée en ce qu'elle appartient à la sous-famille fzf.

5                   2) Protéine selon la revendication 1, caractérisée en ce qu'elle présente au moins 45%, et de préférence au moins 55% d'identité avec le polypeptide de séquence SEQ ID NO:2 ou au moins 60% et de préférence au moins 70 % de similarité avec le polypeptide de séquence  
10                   SEQ ID NO:2.

3) Fragment d'acide nucléique purifié caractérisé en ce qu'il comprend tout ou partie d'une séquence codant pour une protéine selon la revendication 1, ou de sa séquence complémentaire.

15                   4) Vecteur recombinant contenant un fragment d'acide nucléique selon la revendication 3.

5) Cellule transformée par au moins un fragment d'acide nucléique selon la revendication 3.

20                   6) Cellule transformée selon la revendication 5, caractérisée en ce qu'il s'agit d'une cellule végétale.

7) Plante transgénique transformée par au moins un fragment d'acide nucléique selon la revendication 3.

25                   8) Utilisation d'une protéine selon l'une quelconque des revendications 1 ou 2, ou d'une séquence d'acide nucléique selon la revendication 3, pour réguler la différenciation et la prolifération de cellules végétales.

30                   9) Utilisation selon la revendication 8, caractérisée en ce que ladite protéine ou ladite séquence d'acide nucléique est utilisée pour favoriser l'endoploidie dans les cellules d'une plante ou d'un tissu végétal.

35                   10) Utilisation selon la revendication 8, caractérisée en ce que ladite protéine ou ladite séquence



d'acide nucléique est utilisée pour favoriser la régénération in vitro de plantes à partir de cals en culture.

- 5 11) Utilisation d'une protéine de la sous-famille f2r ou d'une séquence d'acide nucléique codant pour tout ou partie de ladite protéine, ou de sa séquence complémentaire, pour réguler la différenciation et la prolifération de cellules végétales.

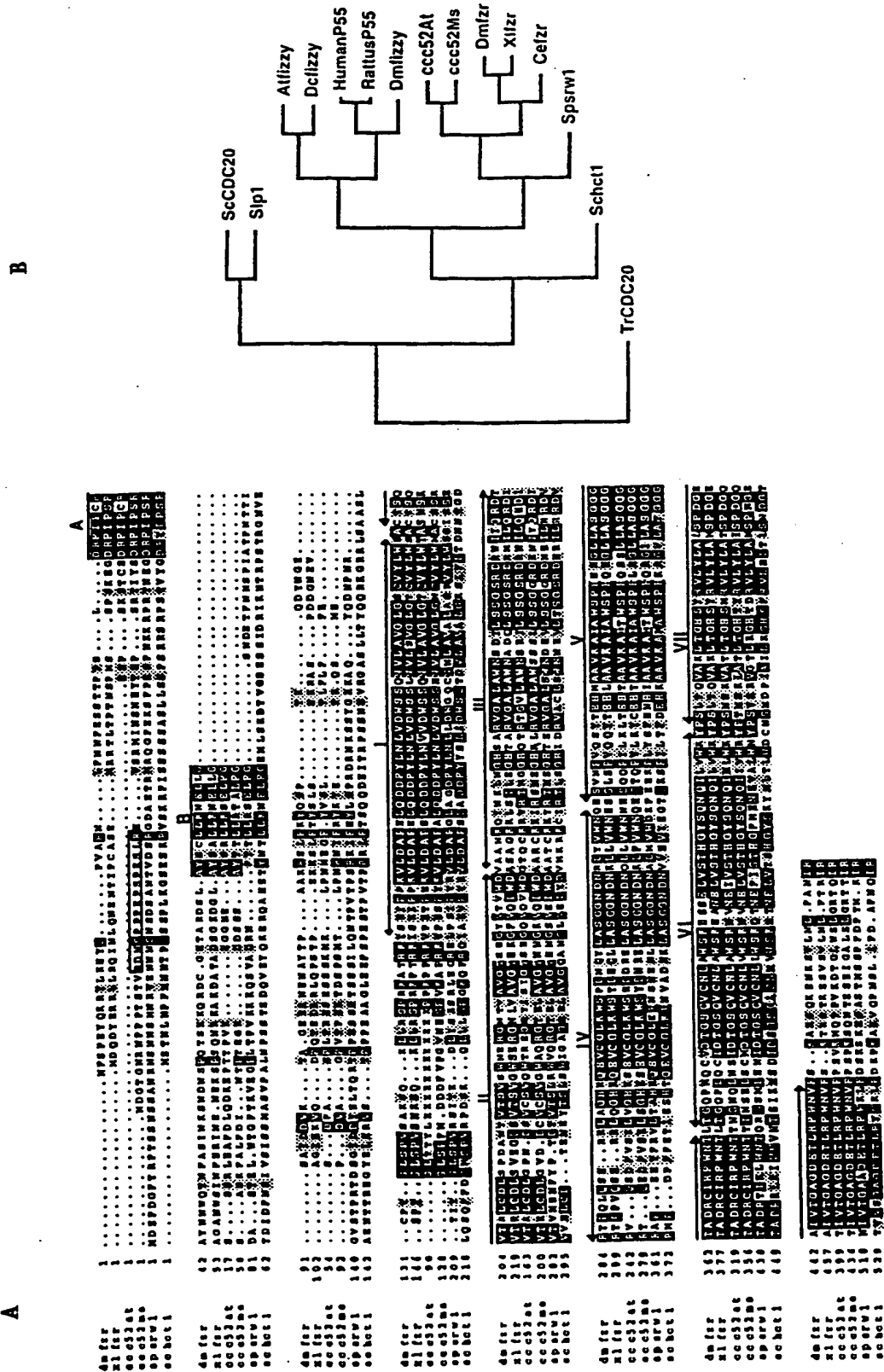


FIGURE 1

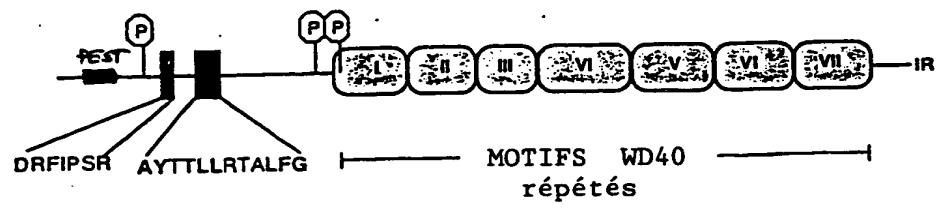
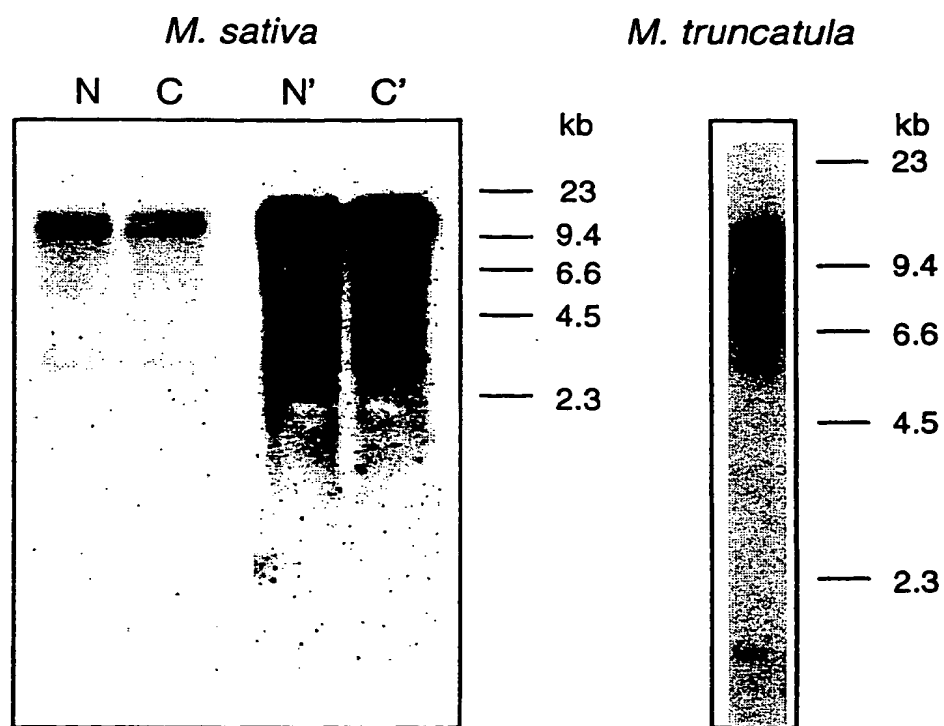


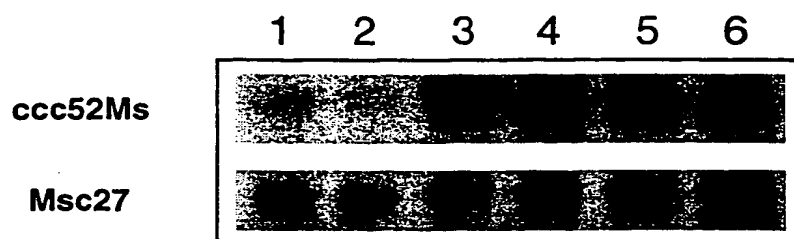
FIGURE 2

3/5

A



B



C

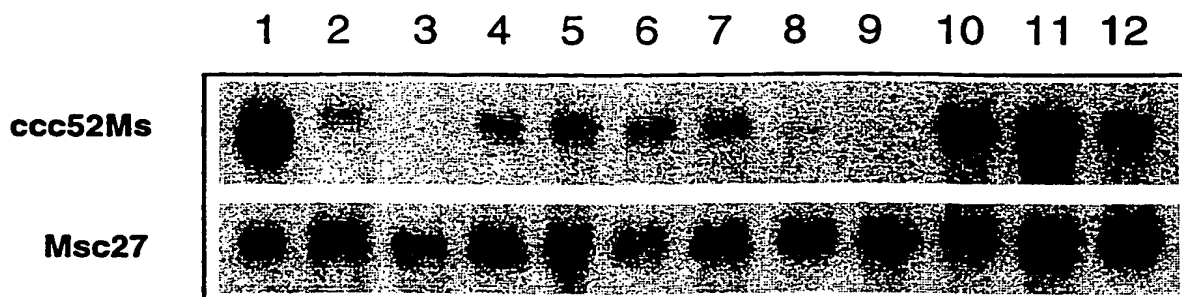


FIGURE 3

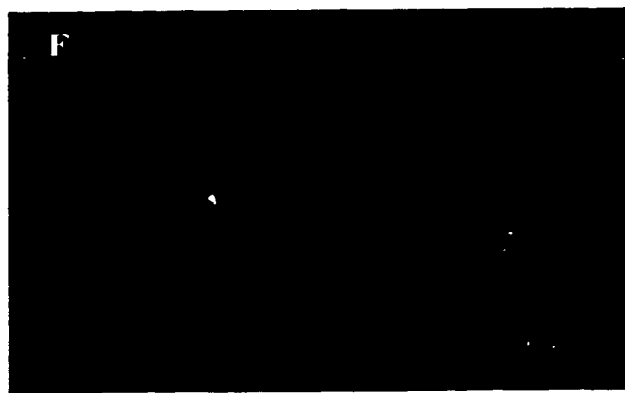
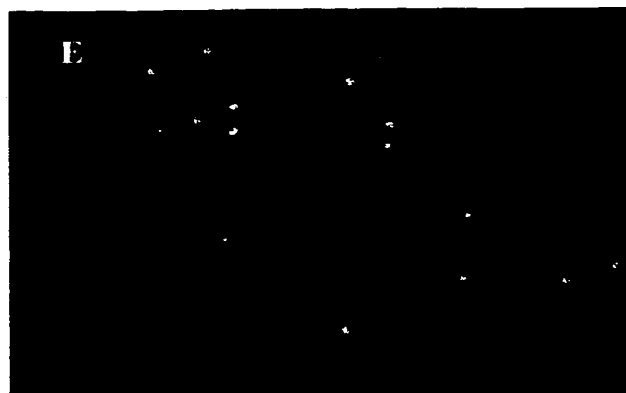
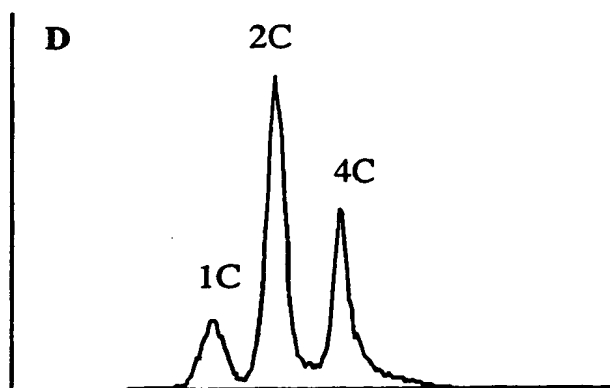
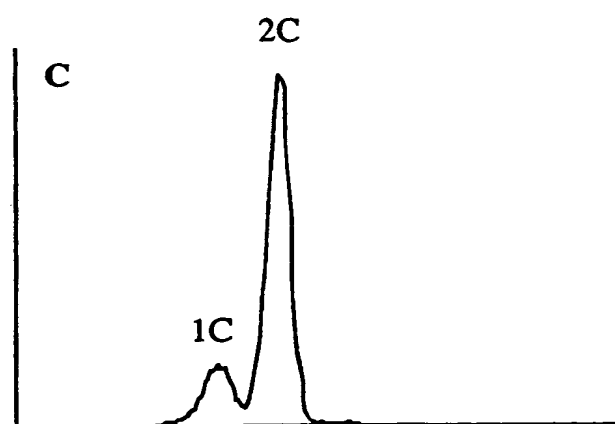
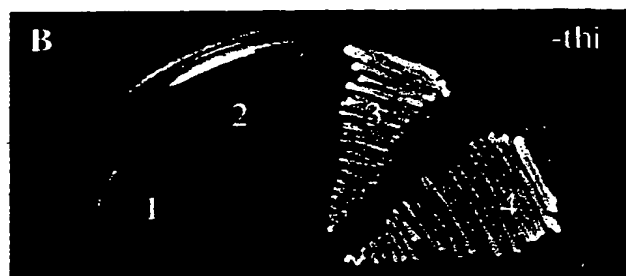


FIGURE 4

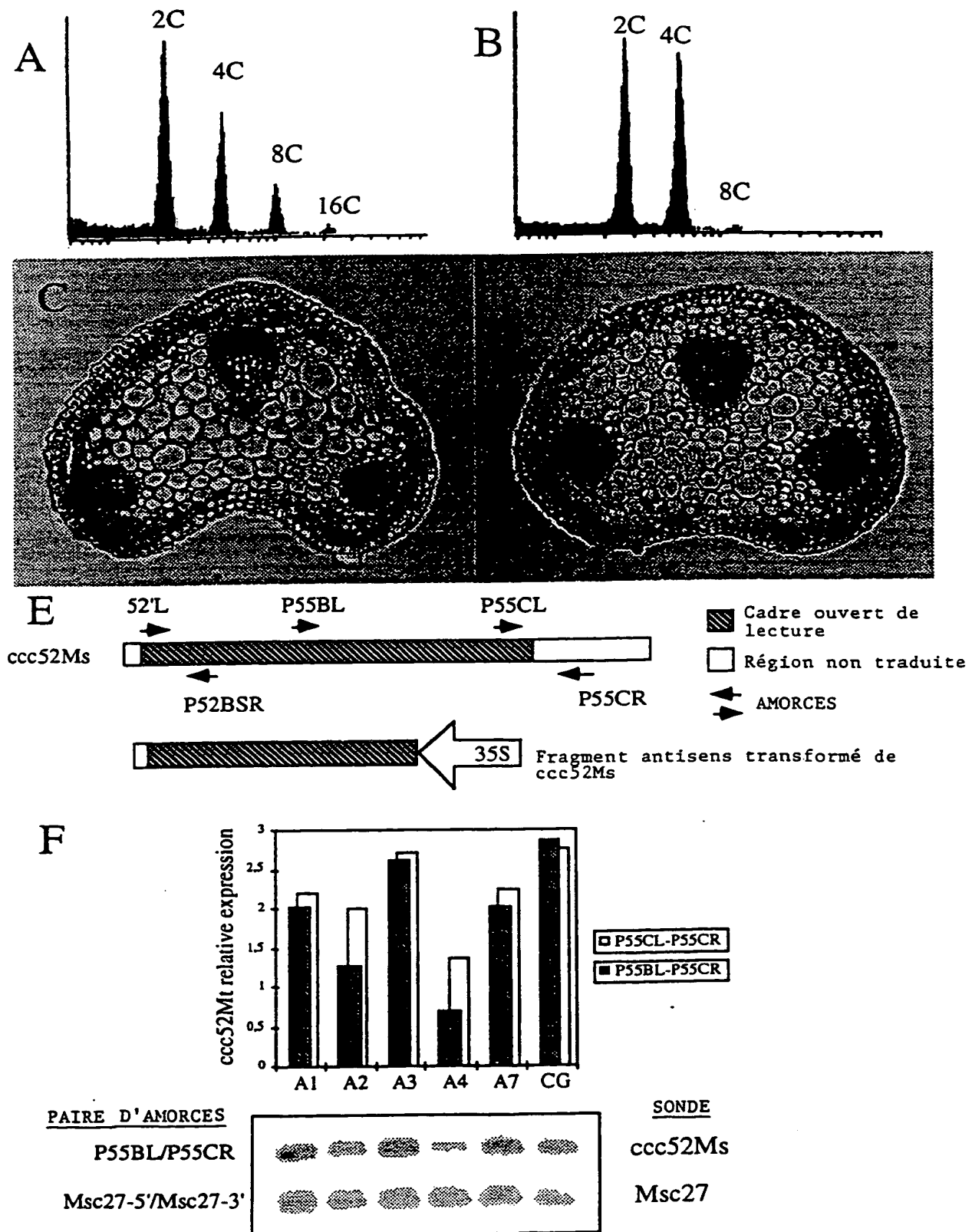


FIGURE 5



1  
2  
3